21/3,K/1 (Item 1 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS

01034848
Apparatus and method for monitoring a computer system to guide optimization
Vorrichtung und Verfahren zur Uberwachung eines Rechnersystems zur Steuerung der Optimierung
Dispositif et procede de surveillance d'un systeme d'ordinateur pour guider son optimisation
PATENT ASSIGNEE:
DIGITAL EQUIPMENT CORPORATION, (313091), 20555 S.H. 249, Houston, Texas 77070-2698, (US), (Applicant designated States: all)
INVENTOR:
Dean, Jeffrey A., 884 Fifteenth Avenue, Menlo Park, California 94025, (US)
Chrysos, George Z., 51 Briarwood Lane, Apt. 5, Marlboro, Massachussetts 01752, (US)
Hicks, James E., 63 Bow Road, Newton, Massachussetts 02159, (US)
Waldspurger, Carl A., 27 Park Drive, Atherton, California 94027, (US)
Weihl, William E., 280 Clipper Street, San Francisco, California 94114, (US)
LEGAL REPRESENTATIVE:
Brunner, Michael John (28871), GILL JENNINGS & EVERY Broadgate House 7 Eldon Street, London EC2M 7LH, (GB)
PATENT (CC, No, Kind, Date): EP 919919 A2 990602 (Basic)
EP 919919 A3 000223
APPLICATION (CC, No, Date): EP 98309634 981125;
PRIORITY (CC, No, Date): US 980124 971126
DESIGNATED STATES: DE; FR; GB
EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO; SI
INTERNATIONAL PATENT CLASS: G06F-011/34
ABSTRACT WORD COUNT: 60
NOTE:
Figure number on first page: 2

LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	9922	473
SPEC A	(English)	9922	7159
Total word count - document A			7632
Total word count - document B			0
Total word count - documents A + B			7632

...SPECIFICATION can be based on some measure of similarity between the instructions, for example, recent branch **history** , stalls, instruction types, or other recent state **history** .

Pinpointing performance bottlenecks in out-of-order processors requires detailed information about **both** stall times and **concurrency levels** . In contrast to in-order processors, a long-latency instruction is not problematic when there is sufficient **concurrency** to efficiently utilize the processor while the long-latency instruction is stalled.

One approach for **obtaining** concurrency information is to snapshot the entire pipeline state. That will directly reveal where sets...compute the average concurrency level when a memory access transaction "hits" in one of the **caches** , and then to compare the average concurrency level with the case where a **cache** miss is incurred. Other interesting aspects to examine for correlation with varying concurrency levels include register dependent stalls, **cache** miss stalls, branch-misprediction stalls, and recent branch **history** .

An additional benefit of profiling a cluster of instructions is the ability to obtain path...

21/3,K/2 (Item 2 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS

00867543

Dynamic set prediction method and apparatus for a multi-level cache system

Verfahren und Gerat zur dynamischen Vorhersage des Weges fur mehrstufige und mehrwege-satz-assoziative Cachespeicher

Methode et appareil pour predire la voie d'une antememoire associative a plusieurs voies et a plusieurs niveaux

PATENT ASSIGNEE:
SUN MICROSYSTEMS, INC., (1392730), 2550 Garcia Avenue, Mountain View, CA 94043, (US), (Proprietor designated states: all)
INVENTOR:
Yung, Robert, 5797 Commerce Drive, Fremont, California 94555, (US)
LEGAL REPRESENTATIVE:
Hogg, Jeffery Keith et al (31905), Withers & Rogers, Goldings House, 2 Hays Lane, London SE1 2HW, (GB)

PATENT (CC, No, Kind, Date): EP 795828 A2 970917 (Basic)
EP 795828 A3 971229
EP 795828 B1 030502
APPLICATION (CC, No, Date): EP 97301126 970221;
PRIORITY (CC, No, Date): US 615662 960313
DESIGNATED STATES: DE; FR; GB
INTERNATIONAL PATENT CLASS: G06F-012/08
ABSTRACT WORD COUNT: 144
NOTE:
Figure number on first page: 8

LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	199709W2	807
CLAIMS B	(English)	200318	459
CLAIMS B	(German)	200318	491
CLAIMS B	(French)	200318	545
SPEC A	(English)	199709W2	4444
SPEC B	(English)	200318	4574
Total word count - document A			5252
Total word count - document B			6069
Total word count - documents A + B			11321

...SPECIFICATION data, it is not necessary to go back and review the last data entry to **predict** the set for the next data entry if this information is stored with the instruction...earlier. The cache 128 is physically direct-mapped, but logically divided into four sets. Each **cache** line includes a data portion 114, a tag portion 112, and level-one set predictor...

...set predictor 116. LRU information 150 is also included. The data (or instruction) from the **cache** is provided to the execution unit of the microprocessor on an I/O bus 118...

...register 132, and the L1 set predictor from the last access in latch 120. The **L1** set predictor for the selected **cache** entry is then provided to latch 120, to override the previous entry and be **ready** for the next **cache** access. At the **same time** , the **level - two** set predictor is provided to a latch 126.

A tag comparator 130 compares the actual input address from register 132 on bus 134 to the tag from the **cache** . This is used to first determine if the appropriate set was properly predicted. If it...

...correctly, a miss signal on line 121 is provided to prefetch unit 122. If the **cache** entry was an instruction, the instruction already loaded into the instruction buffer is invalidated. Where the **cache** entry was data, a miss signal is used to invalidate a data register where the **cache** contents were data loaded into a data register.

Comparator 130 then compares the address to the tags for the other logical sets in the **cache** . This can he done with 3 more accesses (or less, it there is a hit...

this **cache** line, the SP for the previous address can be written at the **same time** . This is possible because there is a separate addressing input for the SP portion of the **L1 cache** , which is addressed by an SP address in a register 137 as shown in Fig. 8.

Alternately, the correction for the SP bits for either the **L1** SP or **L2** SP could be provided to a write-back buffer in **prefetch** unit 122. There, it is stored along with the associated address from previous address register 131, for later writing back to the **L1 cache** upon an empty cycle becoming available. In addition, in a preferred embodiment, the set prediction...

...thereof. For example, the set-prediction information could be stored for each line of a **cache** , or alternately for each entry or for a group of lines. In an alternate embodiment, the invention could be applied to a 3rd level or Nth level **cache** . In addition, the present invention can be applied to other data structures which are not specifically labeled as **caches** . The method set forth above may be varied for different embodiments. For instance, the branch separately for the L1 and L2 **caches** , allowing one to be written while the other is not, or vice versa.

21/3,K/3 (Item 3 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS

00821424
Stream prefetching buffer with stream filter
Datenstromvorausladepufferspeicher mit Datenstromfilters
Memoire tampon pour precharger des flux de donnees et comprenant un filtre de flux
PATENT ASSIGNEE:
International Business Machines Corporation, (200120), Old Orchard Road, Armonk, N.Y. 10504, (US), (Proprietor designated states: all)
INVENTOR:
Mayfield, Michael J., 9111 Glenlake Drive, Austin, Texas 78730, (US)
LEGAL REPRESENTATIVE:
Burt, Roger James, Dr. et al (52152), IBM United Kingdom Limited Intellectual Property Department Hursley Park, Winchester Hampshire SO21 2JN, (GB)
PATENT (CC, No, Kind, Date): EP 763795 A1 970319 (Basic)
EP 763795 B1 011024
APPLICATION (CC, No, Date): EP 96305341 960722;
PRIORITY (CC, No, Date): US 519032 950824
DESIGNATED STATES: DE; FR; GB
INTERNATIONAL PATENT CLASS: G06F-012/08
ABSTRACT WORD COUNT: 205
NOTE:
Figure number on first page: 5

LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	EPAB97	838
CLAIMS B	(English)	200143	433
CLAIMS B	(German)	200143	444
CLAIMS B	(French)	200143	516
SPEC A	(English)	EPAB97	9122
SPEC B	(English)	200143	9446
Total word count - document A			9962
Total word count - document B			10839
Total word count - documents A + B			20801

...SPECIFICATION policy known in the art. The larger L2 cache 203 holds more data than L1 **cache** 202 and ordinarily controls the memory coherency protocol. In the present invention, the data in...stream buffer.

When PCC 404 fetches data, if it is in L1 cache 202 (an **L1** hit), it

is sent to PCC 404. If it is not in **L1** cache 202 (an **L1** miss), but it is in **L2** cache 203 (an **L2** hit), a line of **L1** cache 202 is replaced with this subject data from **L2** cache 203. In this case, the data is sent **simultaneously** to **L1** cache 202 and PCC 404. If there is a miss in **L2** cache 203 as well, the data may be **fetched** from memory 209 into BIU 401 and loaded **simultaneously** into **L1** cache 202, **L2** cache 203, and PCC 404. Variations on this operation are known in the art. Data store operations are similar to the **fetch** operations except that the data is stored into an L1 line to complete the operation...

...miss. The filter contains a number of locations that can hold such addresses comprising a " **history** " of such events. They may be replaced on a least recently used (LRU) basis. Whenever...

...SPECIFICATION policy known in the art. The larger L2 cache 203 holds more data than L1 **cache** 202 and ordinarily controls the memory coherency protocol. In the present invention, the data in...

...stream buffer.
When PCC 404 fetches data, if it is in L1 cache 202 (an **L1** hit), it is sent to PCC 404. If it is not in **L1** cache 202 (an **L1** miss), but it is in **L2** cache 203 (an **L2** hit), a line of **L1** cache 202 is replaced with this subject data from **L2** cache 203. In this case, the data is sent **simultaneously** to **L1** cache 202 and PCC 404. If there is a miss in **L2** cache 203 as well, the data may be **fetched** from memory 209 into BIU 401 and loaded **simultaneously** into **L1** cache 202, **L2** cache 203, and PCC 404. Variations on this operation are known in the art. Data store operations are similar to the **fetch** operations except that the data is stored into an L1 line to complete the operation...

...miss. The filter contains a number of locations that can hold such addresses comprising a " **history** " of such events.

21/3,K/4 (Item 4 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS

00819598
Prefetching data cache
Cachespeicher mit Datenvorausladung
Antememoire avec prechargement de donnees
PATENT ASSIGNEE:
International Business Machines Corporation, (200120), Old Orchard Road, Armonk, N.Y. 10504, (US), (Proprietor designated states: all)
INVENTOR:
Mayfield, Michael J., 9111 Glenlake Drive, Austin, Texas 78730, (US)
LEGAL REPRESENTATIVE:
Burt, Roger James, Dr. et al (52152), IBM United Kingdom Limited Intellectual Property Department Hursley Park, Winchester Hampshire SO21 2JN, (GB)
PATENT (CC, No, Kind, Date): EP 762288 A2 970312 (Basic)
EP 762288 A3 970319
EP 762288 B1 011031
APPLICATION (CC, No, Date): EP 96305419 960724;
PRIORITY (CC, No, Date): US 519031 950824
DESIGNATED STATES: DE; FR; GB
INTERNATIONAL PATENT CLASS: G06F-012/08
ABSTRACT WORD COUNT: 76
NOTE:
Figure number on first page: 4

LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:
Available Text Language Update Word Count
CLAIMS A (English) EPAB97 1188
CLAIMS B (English) 200144 645
CLAIMS B (German) 200144 676

CLAIMS B	(French)	200144	773
SPEC A	(English)	EPAB97	5993
SPEC B	(English)	200144	6085
Total word count - document A			7182
Total word count - document B			8179
Total word count - documents A + B			15361

...SPECIFICATION policy known in the art. The larger L2 cache 203 holds more data than L1 **cache** 202 and ordinarily controls the memory coherency protocol. In the present invention, the data in...

...stream buffer.
When PCC 404 fetches data, if it is in L1 cache 202 (an **L1** hit), it is sent to PCC 404. If it is not in **L1** cache 202 (an **L1** miss), but it is in **L2** cache 203 (an **L2** hit), a line of **L1** cache 202 is replaced with this subject data from **L2** cache 203. In this case, the data is sent **simultaneously** to **L1** cache 202 and PCC 404. If there is a miss in **L2** cache 203 as well, the data may be **fetched** from memory 209 into BIU 401 and loaded **simultaneously** into **L1** cache 202, **L2** cache 203, and PCC 404. Variations on this operation are known in the art. Data store operations are similar to the **fetch** operations except that the data is stored into an L1 line to complete the operation...

...miss. The filter contains a number of locations that can hold such addresses comprising a " **history** " of such events. They may be replaced on a least recently used (LRU) basis. Whenever...

...SPECIFICATION policy known in the art. The larger L2 cache 203 holds more data than L1 **cache** 202 and ordinarily controls the memory coherency protocol. In the present invention, the data in...stream buffer.
When PCC 404 fetches data, if it is in L1 cache 202 (an **L1** hit), it is sent to PCC 404. If it is not in **L1** cache 202 (an **L1** miss), but it is in **L2** cache 203 (an **L2** hit), a line of **L1** cache 202 is replaced with this subject data from **L2** cache 203. In this case, the data is sent **simultaneously** to **L1** cache 202 and PCC 404. If there is a miss in **L2** cache 203 as well, the data may be **fetched** from memory 209 into BIU 401 and loaded **simultaneously** into **L1** cache 202, **L2** cache 203, and PCC 404. Variations on this operation are known in the art. Data store operations are similar to the **fetch** operations except that the data is stored into an L1 line to complete the operation...

...miss. The filter contains a number of locations that can hold such addresses comprising a " **history** " of such events. They may be replaced on a least recently used (LRU) basis. Whenever...

21/3,K/5 (Item 5 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS

00345742
Improved multiprocessor cache and method for maintaining coherence
Mehrprozessor-Cachespeicher und Verfahren zur Aufrechterhaltung der Cache-Koherenz
Antememoire de multiprocesseur et methode pour maintenir coherence
PATENT ASSIGNEE:
International Business Machines Corporation, (200120), Old Orchard Road, Armonk, N.Y. 10504, (US), (applicant designated states: DE;FR;GB)
INVENTOR:
Liu, Lishing, 196 Deerfield Lane North, Pleasantville New York 10570, (US)
LEGAL REPRESENTATIVE:
Schafer, Wolfgang, Dipl.-Ing. (62021), IBM Deutschland Informationssysteme GmbH Patentwesen und Urheberrecht, 70548 Stuttgart, (DE)
PATENT (CC, No, Kind, Date): EP 355320 A2 900228 (Basic)
EP 355320 A3 910710

APPLICATION (CC, No, Date): EP 89111606 890626;
PRIORITY (CC, No, Date): US 232722 880816
DESIGNATED STATES: DE; FR; GB
INTERNATIONAL PATENT CLASS: G06F-012/08;
ABSTRACT WORD COUNT: 101

LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	EPABF1	670
CLAIMS B	(English)	EPAB97	1414
CLAIMS B	(German)	EPAB97	1267
CLAIMS B	(French)	EPAB97	1519
SPEC A	(English)	EPABF1	5090
SPEC B	(English)	EPAB97	5103
Total word count - document A			5760
Total word count - document B			9303
Total word count - documents A + B			15063

...SPECIFICATION R bits on DLAT hits/misses. The present invention is independent of the second level **cache** (L2) hierarchy and is only concerned about the concurrency of data lines in a first...

...Patent 4,181,937 to Hattori et al. an MP cache replacement scheme in a **two level** cache hierarchy is taught. Upon the decision of replacement of a block from **L2** shared by all processors, blocks with fewer numbers of copies in the **first level** processor caches are given higher preference. This is supposed to increase **concurrency** at **L1** with better **L2** replacement strategies. The present invention is not concerned with **L2** replacements.
In U.S. Patent 4,503,497 to Krygowski et al. a cache to...

...EX (not CH) only when it is found CH'd in the remote cache. The **cache** to **cache** transfer environment is not discussed in Flusche et al. The present invention provides a capability...

...do this kind of EX (but also CH) fetch upon remote CH situations for a **cache** to **cache** transfer environment. This has several advantages since when using a **cache** to **cache** transfer facility the CH line may be transferred to another cache as EX and CH...restriction on concurrency of a cache line due to its being modified as a remote **past** event, and allows cache lines to be read only and be available to all processors...

...SPECIFICATION R bits on DLAT hits/misses. The present invention is independent of the second level **cache** (L2) hierarchy and is only concerned about the concurrency of data lines in a first...

...in the present invention is achieved by allowing a line in different caches in a **read** only state when a damaging sharing characteristic disappears.
In U.S. Patent 4,181,937 to Hattori et al. an MP cache replacement scheme in a **two level** cache hierarchy is taught. Upon the decision of replacement of a block from **L2** shared by all processors, blocks with fewer numbers of copies in the **first level** processor caches are given higher preference. This is supposed to increase **concurrency** at **L1** with better **L2** replacement strategies. The present invention is not concerned with **L2** replacements.
In U.S. Patent 4,503,497 to Krygowski et al. a cache to...

...CH) only when it is found CH'd in the remote cache. The cache to **cache** transfer environment is not discussed in Flusche et al. The present invention provides a capability...

...do this kind of EX (but also CH) fetch upon remote CH situations for a **cache** to **cache** transfer environment. This has several advantages since when using a **cache** to **cache** transfer facility the CH line may be transferred to another cache as EX and CH...restriction on concurrency of

a cache line due to its being modified as a remote **past** event, and allows cache lines to be read only and be available to all processors...

21/3,K/6 (Item 1 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT

00987019 **Image available**
BRANCH PREDICTION DEVICE WITH TWO LEVELS OF BRANCH PREDICTION CACHE
DISPOSITIF DE PREDICTION DE BRANCHEMENT AVEC CACHE DE PREDICTION DE BRANCHEMENT A DEUX NIVEAUX
Patent Applicant/Assignee:
ADVANCED MICRO DEVICES INC, One AMD Place, Mail Stop 68, Sunnyvale, CA 94088-3453, US, US (Residence), US (Nationality), (For all designated states except: US)
Patent Applicant/Inventor:
ZURASKI Gerald D, 8009 Mead Parke Cv, Austin, TX 78726, US, US (Residence), US (Nationality), (Designated only for: US)
ROBERTS James S, 1621 Partridge Court, #2, Arlington Heights, IL 60004, US, US (Residence), US (Nationality), (Designated only for: US)
Legal Representative:
DRAKE Paul S (agent), Advanced Micro Devices, Inc., 5204 East Ben White Boulevard, Mail Stop 562, Austin, TX 78741 (et al), US,
Patent and Priority Information (Country, Number, Date):
Patent: WO 200317091 A1 20030227 (WO 0317091)
Application: WO 2002US20481 20020627 (PCT/WO US0220481)
Priority Application: US 2001912011 20010724
Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZM ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Publication Language: English
Filing Language: English
Fulltext Word Count: 16478

Fulltext Availability:
Detailed Description

Detailed Description
... storage entry is not occupied by a valid entry, an L2 predictor storage 260 is **queried** for a branch prediction entry corresponding to the **fetch** address (block 408). In one embodiment, Ll predictor storage 206 and **L2** predictor storage 260 may be **queried** in **parallel** . If no corresponding entry is present in the **L2** predictor storage 260 (block 41 0), a new branch prediction entry may be created in the **L 1** predictor storage 206 for the presented **fetch** address (block 412). On the other hand, if there exists an entry in the **L2** branch predictor storage 260 corresponding to the **fetch** address, data from the L2 entry is utilized to rebuild a fall branch prediction corresponding to the **fetch** address (block 414). The rebuilt branch prediction is then stored in the Ll branch predictor...

...recovered from an L2 branch predictor storage, rather than having to rebuild it through a **history** of branch executions. Further, only a subset of information corresponding ...one embodiment, local predictor storage 206 may be organized in the same manner as instruction **cache** 16. Data stored in local predictor storage 206 may consist of lines of storage organized...

...local predictor storage 206 is of sufficient size to cover all entries in the instruction **cache** 16. In an alternative embodiment, local predictor storage 206 may be smaller than instruction **cache** 16. For example, local predictor storage 206 may be 1/4 the size of instruction

cache 16. In such an embodiment, additional bits may be stored along with a local prediction...

21/3,K/7 (Item 2 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2004 WIPO/Univentio. All rts. reserv.

00551255 **Image available**
A METHOD AND APPARATUS FOR BRANCH PREDICTION USING A SECOND LEVEL BRANCH PREDICTION TABLE
PROCEDE ET DISPOSITIF PERMETTANT LES PREVISIONS DE BRANCHE AU MOYEN D'UNE TABLE DE PREVISION DE BRANCHE DE SECOND NIVEAU
Patent Applicant/Assignee:
INTEL CORPORATION,
YEH Tse-Yu,
SHARANGPANI Harshvardhan P,
Inventor(s):
YEH Tse-Yu,
SHARANGPANI Harshvardhan P,
Patent and Priority Information (Country, Number, Date):
Patent: WO 200014628 A1 20000316 (WO 0014628)
Application: WO 99US19892 19990826 (PCT/WO US9919892)
Priority Application: US 98149885 19980908
Designated States: AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG US UZ VN YU ZA ZW GH GM KE LS MW SD SL SZ UG ZW AM AZ BY KG KZ MD RU TJ TM AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
Publication Language: English
Fulltext Word Count: 7483

Fulltext Availability:
Claims

Claim
... determined to be a branch instruction.
. The method of claim 15, wherein the steps of **searching** the **first level** BPT and **searching** the **second level** BPT occur **simultaneously** .

19 The method of claim 15, wherein the step of **searching** the **first level** BIPT includes the step of comparing an address tag of the IP address to an address tag stored in the **first level** BPT, and the step of **searching** the **second level** BPT includes the step of selecting an entry from a directmapped table.
20 A method...

...of:
predicting the subsequent IP address to be a target address from a target address **cache** if a branch prediction entry in a first branch prediction table (BPT) associated with the...branch is not taken.

22 The method of claim 20, further comprising the step of **predicting** the subsequent IP address to be the initial IP address incremented by a predetermined amount...

21/3,K/8 (Item 3 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2004 WIPO/Univentio. All rts. reserv.

00196049 **Image available**

TWO-LEVEL BRANCH PREDICTION CACHE

ANTEMEMOIRE DE PREDICTION DE BRANCHEMENT A DEUX NIVEAUX

Patent Applicant/Assignee:
NEXGEN MICROSYSTEMS,
Inventor(s):
STILES David R,
FAVOR John G,
VAN DYKE Korbin S,

Patent and Priority Information (Country, Number, Date):
Patent: WO 9113402 A1 19910905
Application: WO 91US826 19910206 (PCT/WO US9100826)
Priority Application: US 90306 19900226
Designated States: AT BE CH DE DE DK ES FR GB GB GR IT JP KR LU NL SE
Publication Language: English
Fulltext Word Count: 9090

Fulltext Availability:
Detailed Description

Detailed Description

... this invention, a second level cache entry holds only a partial target address and one **history** bit. The **predicted** direction of a conditional branch is based simply on the direction last taken by that...

...within a subset of the instruction address space also containing the branch instruction. The full **predicted** target ...incorporating the present invention; Fig. 2 is an overall block diagram of the branch prediction **cache** (BPC) and its immediate environment;
Fig. 3 is a block diagram of the first level...

...elements within CPU 10, not such external devices.

An Instruction Decoder (DEC) 12 performs instruction **fetch** , instruction decode, and pipeline control. DEC 12 optionally interleaves instruction **prefetch** of up to three **simultaneous** instruction streams. DEC 12 contains a **two - level** Branch Prediction **Cache** (BPC) 13. Ile BPC includes an integrated structure which contains dynamic branch **history** data, a physical branch target address, and a branch target buffer for each **cache** entry. As branch instructions are decoded, the BPC is consulted for information about that branch. Independent of the direction **predicted** , branches are executed in a single cycle and do not cause pipeline bubbles.

On each...in greater detail below, first level BPC 152 is a shallow but wide structure which **caches** full prediction information for a limited number of branch instructions. In particular, first level BPC...

...instruction valid (TIV) bits. Second level BPC 155 is a deep but narrow structure which **caches** only partial prediction information but for a much larger number of branch instructions. Second level...

...entries, each containing two bytes of partial target address information and one history bit.

In **parallel** with instruction decoding, the instruction's decode PC is used to perform **parallel** **lookups** in the first and **second** **level** BPCs. (Since the incoming instructions have not been decoded at this point, non-branch instructions are also **checked**). In the event of a hit on **first** **level** BPC 152, the target instruction bytes are communicated to instruction decoder 160, the branch history...

...BPC 155 is always assumed to hit. Therefore, second level BPC 155 communicates the branch **history** bit to IDC 162 and the partial target address to IFC 165 on every access...

...fetch address is offset from the target address by the number of target

instruction bytes **cached** for that branch. In the event of a miss, the fetch address and the target address are the same.

Note that prediction information for a branch, i.e., a valid **cache** entry associated with the instruction, is created only after a branch is encountered at least once and continues to exist in the **cache** only until replaced by a set of prediction information for another branch. As with most **caching** strategies, the benefit of these **cache** schemes primarily exists on a statistical basis. When the cache does not contain an entry...mention above, a second level BPC entry holds only a partial target address and one **history** bit. The predicted direction of a conditional branch is based simply on the direction last...the number of cache entries.

Operation
A first level cache size of 36 entries and **second level** cache size of 256 entries, in combination with a factor of 16 difference in per-entry cost, results in second-to- **first level** cache ratios of eight times the number of entries, yet still almost half the size. With this much larger size, even given the direct-mapped organization, the **second level** cache provides an effective backup to the **first level** cache.

As each branch instruction is fetched its address is used to perform parallel look-ups in the two levels of BPC: the large-set or fully associative **first level** access using the full branch address; and the direct-mapped, tag-less **second level** using only a subset of the address bits for the index.

If there is a tag match with a **first level** cache entry, then all of this entry's prediction information is read out, and the **second** level BPC is ignored. All the necessary predictions are made, effectively eliminating or hiding any...still be a delay before processing of target instructions can begin, if the branch is **predicted** taken.